WHAT IS CLAIMED IS:

1. A method for sampling assets in an asset portfolio for optimal underwriting coverage when only a portion of the assets are to be underwritten, said method comprising the steps of:

determining descriptive attributes of assets in the portfolio;

encoding individual attributes; and

clustering the assets for underwriting based upon occurrences of the descriptive attributes.

2. A method according to Claim 1 further comprising the steps of determining a number of samples to be submitted for further underwriting review.

3. A method according to Claim 2 wherein said step of determining a number of samples to be submitted for further underwriting review further comprises the steps of:

establishing a confidence level regarding the total recoveries probable in each segment of the portfolio;

establishing a precision to which total recoveries in each segment are estimated; and

providing an estimate of a level and a range of recoveries as a percentage of total Unpaid Principal Balance (UPB).

4. A method according to Claim 3 wherein said step of establishing a confidence level regarding the total recoveries probable further comprises the step of determining a sample size, n, for the cluster of assets according to:

$$h^2 = k^2 \times n\left[1 - \frac{n}{N}\right] \times \frac{\left[\sum_1^N x_i\right]^2}{\left[\sum_1^n x_i\right]^2} \times \frac{\sum_1^N (y_i - Rx_i)^2}{N-1}$$

h = *desired precision*

n = sample size

N = cluster size

x_i = UPB for sample i

y_i = recovery for sample i

$R = \dfrac{\sum_1^N y_i}{\sum_1^N x_i}$ = cluster expected recovery %

h = error tolerance for estimating $Y = \sum_1^N y_i$ with $\hat{Y}_R$

and solving for n.

5. A method according to Claim 4 wherein said step of providing an estimate of a level and a range of recoveries further comprises the step of estimating a level and range of recoveries according to:

$$\hat{Y}_R = \hat{R} \times \sum_{i=1}^N x_i = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i} \times \sum_{i=1}^N x_i = \frac{\sum_{i=1}^n \rho_i x_i}{\sum_{i=1}^n x_i} \times \sum_{i=1}^N x_i$$

k = constant in Tchebyshev's Formula :

$$\left|\hat{Y}_R - \mu_{\hat{Y}_R}\right| \le k\sqrt{Var(\hat{Y}_R)} \text{ with probability} \ge 1 - \frac{1}{k^2}$$

6. A method according to Claim 1 wherein said step of clustering the assets for underwriting further comprises the step of using a supervised clustering process to cluster the assets.

7. A method according to Claim 1 wherein said step of clustering the assets for underwriting further comprises the step of using an unsupervised clustering process to cluster the assets.

8. A method according to Claim 1 wherein said step of clustering the assets for underwriting further comprises the step of using a Monte Carlo process to cluster the assets.

9. A system configured to sample assets in an asset portfolio for optimal underwriting coverage, said system comprising:

a computer configured as a server and further configured with a database of asset portfolios and to enable valuation process analytics;

at least one client system connected to said server through a network, said server further configured to:

determine descriptive attributes of assets in the portfolio;

encode individual attributes; and

cluster the assets for underwriting based upon occurrences of the descriptive attributes.

10. A system according to Claim 9 further configured to determine a number of samples to be submitted for further underwriting review.

11. A system according to Claim 10 wherein said server configured to:

establish a confidence level regarding the total recoveries probable in each segment of the portfolio;

establish a precision to which total recoveries in each segment are estimated; and

provide an estimate of a level and a range of recoveries as a percentage of total Unpaid Principal Balance (UPB).

12. A system according to Claim 11 wherein said server configured to determine a sample size, n, for the cluster of assets according to:

$$h^2 = k^2 \times n\left[1 - \frac{n}{N}\right] \times \frac{\left[\sum_1^N x_i\right]^2}{\left[\sum_1^n x_i\right]^2} \times \frac{\sum_1^N (y_i - Rx_i)^2}{N-1}$$

h = *desired precision*

n = sample size

N = cluster size

x_i = UPB for sample i

y_i = recovery for sample i

$$R = \frac{\sum_1^N y_i}{\sum_1^N x_i} = \text{cluster expected recovery \%}$$

h = error tolerance for estimating $Y = \sum_1^N y_i$ with $\hat{Y}_R$

by solving for n.

13. A system according to Claim 12 wherein said server configured to estimate a level and range of recoveries according to:

$$\hat{Y}_R = \hat{R} \times \sum_{i=1}^N x_i = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i} \times \sum_{i=1}^N x_i = \frac{\sum_{i=1}^n \rho_i x_i}{\sum_{i=1}^n x_i} \times \sum_{i=1}^N x_i$$

k = constant in Tchebyshev's Formula :

$$\left|\hat{Y}_R - \mu_{\hat{Y}_R}\right| \le k\sqrt{Var(\hat{Y}_R)} \text{ with probability} \ge 1 - \frac{1}{k^2}.$$

14. A system according to Claim 9 wherein said server configured to use a supervised clustering process to cluster the assets.

15. A system according to Claim 9 wherein said server configured to use an unsupervised clustering process to cluster the assets.

16. A system according to Claim 9 wherein said server configured to use a Monte Carlo process to cluster the assets.

17. A computer for sampling assets in an asset portfolio for optimal underwriting coverage, said computer including a database of asset portfolios and valuation process analytics, said computer programmed to:

determine descriptive attributes of assets in the portfolio;

encode individual attributes; and

cluster the assets for underwriting based upon occurrences of the descriptive attributes.

18. A computer according to Claim 17 programmed to determine a number of samples to be submitted for further underwriting review.

19. A computer according to Claim 18 programmed to:

establish a confidence level regarding total recoveries probable in each segment of the portfolio;

establish a precision to which total recoveries in each segment are estimated; and

provide an estimate of a level and a range of recoveries as a percentage of total Unpaid Principal Balance (UPB).

20. A computer according to Claim 19 programmed to determine a sample size, n, for the cluster of assets according to:

$$h^2 = k^2 \times n\left[1 - \frac{n}{N}\right] \times \frac{\left[\sum_1^N x_i\right]^2}{\left[\sum_1^n x_i\right]^2} \times \frac{\sum_1^N (y_i - Rx_i)^2}{N-1}$$

h = *desired precision*

n = sample size

N = cluster size

x_i = UPB for sample i

y_i = recovery for sample i

$$R = \frac{\sum_1^N y_i}{\sum_1^N x_i} = \text{cluster expected recovery \%}$$

h = error tolerance for estimating $Y = \sum_1^N y_i$ with $\hat{Y}_R$

by solving for n.

21. A computer according to Claim 20 programmed to estimate a level and range of recoveries according to:

$$\hat{Y}_R = \hat{R} \times \sum_{i=1}^N x_i = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i} \times \sum_{i=1}^N x_i = \frac{\sum_{i=1}^n \rho_i x_i}{\sum_{i=1}^n x_i} \times \sum_{i=1}^N x_i$$

k = constant in Tchebyshev's Formula :.

$$\left|\hat{Y}_R - \mu_{\hat{Y}_R}\right| \le k\sqrt{Var(\hat{Y}_R)} \text{ with probability} \ge 1 - \frac{1}{k^2}.$$

22. A computer according to Claim 17 programmed to use a supervised clustering process to cluster the assets.

23. A computer according to Claim 17 programmed to use an unsupervised clustering process to cluster the assets.

24. A computer according to Claim 17 programmed to use a Monte Carlo process to cluster the assets.

09376280-121500